

# *Estimating wind speed and capacity factors in Mexico using reanalysis data*

Article

Accepted Version

Creative Commons: Attribution-Noncommercial-No Derivative Works 4.0

Morales-Ruvalcaba, C. F., Rodriguez-Hernandez, O., Martinez-Alvarado, O. ORCID: <https://orcid.org/0000-0002-5285-0379>, Drew, D. R. and Ramos, E. (2020) Estimating wind speed and capacity factors in Mexico using reanalysis data. *Energy for Sustainable Development*, 58. pp. 158-166. ISSN 0973-0826 doi: <https://doi.org/10.1016/j.esd.2020.08.006> Available at <https://centaur.reading.ac.uk/92440/>

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

To link to this article DOI: <http://dx.doi.org/10.1016/j.esd.2020.08.006>

Publisher: Elsevier

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the [End User Agreement](#).

[www.reading.ac.uk/centaur](http://www.reading.ac.uk/centaur)

**CentAUR**

Central Archive at the University of Reading

Reading's research outputs online

# Estimating wind speed and capacity factors in Mexico using reanalysis data

C. F. Morales-Ruvalcaba<sup>1</sup>, O. Rodriguez-Hernandez<sup>2</sup>,  
O. Martínez-Alvarado<sup>3</sup>, D.R. Drew<sup>4</sup>, E. Ramos<sup>2</sup>

August 22, 2020

<sup>1</sup> Licenciatura en Ingeniería en Energías Renovables, Instituto de Energías Renovables. Universidad Nacional Autónoma de México, A.P. 34, 62580 Temixco, Mor. México, Mexico

<sup>2</sup> Instituto de Energías Renovables. Universidad Nacional Autónoma de México, A.P. 34, 62580 Temixco, Mor. México, Mexico

<sup>3</sup> National Centre for Atmospheric Science, Department of Meteorology, University of Reading, Earley Gate, RG6 6BB, United Kingdom

<sup>4</sup> Department of Meteorology, University of Reading, Earley Gate, RG6 6BB, United Kingdom

## Abstract

In the fundamental stage of resource assessment, high-quality wind speed measurements are required to estimate power production. However, this high-quality data is not always available, and therefore the analysis of alternative sources becomes essential. In this work, we analyze the ability of MERRA-2 to represent wind speed characteristics at 24 anemometric stations in Mexico. The assessment was carried out using the Pearson correlation coefficient between the observed time series, and the obtained by interpolating bias-corrected reanalysis-estimated wind speed to all locations for different time-averaging periods. Results showed that the reanalysis' performance is not uniform throughout the country; it depends on the time resolution, local orographic conditions, and the relationship between the local flow and the large-scale circulation. Based on these results, the country was subdivided into eight regions. The best-represented region was the Chivela Pass, where the winds are tightly linked to the interaction between the large-scale circulation and the local orography. The worst performing regions were located where the land sea-mask and orography at the reanalysis' resolution may not be accurate enough to reproduce the station's wind speeds. Reanalysis-estimated capacity factors exhibit large interannual variability in some stations, which can have significant consequences for the operation of individual wind farms and the power grid. The results show that, while caution should be exercised when applying reanalyses to wind resource assessment in Mexico, reanalysis wind power estimates can be a valuable tool to investigate the feasibility and installed capacity requirements for Mexico to meet its renewable energy targets.



# 1 Introduction

The potential impacts of climate change have encouraged actions to mitigate greenhouse gas emissions from fossil fuel combustion. Renewable sources of energy play a critical role in this context, in particular, solar photovoltaic and wind power technologies due to their technical maturity and competitive prices. Globally, wind power is a popular alternative for clean electric generation. In 2018, 52.4 GW of capacity was installed, taking the global cumulative capacity to 597 GW. This growth looks set to continue with some studies forecasting a capacity of 817 GW by 2021 [16].

Mexico, among others countries, is considered a key market for renewable energy. As of 2017, renewable energy contributed 26% of total electricity generation in Mexico, including hydro-electric (11%) and geothermal. Wind power currently contributes 3% of the electric power generation [2]; however, the capacity is increasing. In 2017, Mexico installed 478 MW of wind power capacity resulting in a record low price of US\$0.017/kWh for one project [16]. This growth looks set to continue as the Mexican government has set a series of ambitious renewable energy targets, including increasing the installed capacity of wind power to 15.1 GW by 2030 [1].

Although some wind rich sites are well identified in Mexico [20], there is little to no information of the wind resources in the rest of the country. A number of studies have investigated the wind resource for specific regions in Mexico [17] [19] [13] [6] [18]. Resource assessments in Mexico are mainly estimated by using the meteorological weather stations service, where 10-m winds measured at meteorological surface stations are used to feed the assessment models [32]. There are only 186 locations for which hourly observations are collected, which means large areas of the vast Mexican territory of nearly 2 million km<sup>2</sup> have few observations. Furthermore many of the stations are relatively new and data is only available for a short period. Given that the wind exhibits variability at all time scales from sub-minute to seasonal, it is important to have wind records as frequently as possible and for long periods of time, preferably several years.

To address the problem of a lack of long-term wind speed data, a number

of studies have used reanalysis products to estimate wind power characteristics. For example, NASA’s Modern-Era Retrospective Analysis for Research and Applications (MERRA) [28] has been used to calculate the nationwide wind energy production in Great Britain [5], Sweden [25], and Ireland [9], as well as for large regions in the US Great Plains [29] and Southern Canada [11]. Other studies have estimated wind energy production from NASA’s MERRA version 2 (MERRA-2) [15] to analyse Saudi Arabia [7], India [14], Central Romania [3], California [40, 23], the US Northwest, France, Germany, Denmark, Sweden [24], and Europe [8, 38]. MERRA-2 is also used for offshore wind resource assessment in South Korea [21], the Eastern US Coast [22] and different countries around the globe [4]. However, none of the previous mentioned works focuses on Mexico. The country’s location, between 15°N and 32°N and between the Eastern Pacific and the North Atlantic, leads to its weather and climate being influenced by both tropical (e.g. tropical cyclones from both ocean basins) and extratropical weather systems (e.g. fronts associated with the passage of extratropical cyclones at mid-latitudes). This, together with its complex topography, still represents a challenge even for state-of-the-art global atmospheric models due to limitations related to model resolution and the parametrisation of sub-grid surface processes.

There are clear advantages of using reanalyses as data sources such as its global coverage, the information is accessible for anyone, and it is available over large periods of time. However there are disadvantages like its spatial-temporal resolution is still low when compared with other sources of data, and it is possible that very specific characteristics of one site might not be well represented by the reanalysis. According to this, the very mountainous orography of the Mexican territory, as well as its large coastline may affect the reanalysis’ behaviour. The aim of this study is to investigate how well the MERRA-2 reanalysis data represents the wind resource across Mexico and consequently determine its applicability for assessing wind power characteristics.

The rest of the article is organised as follows. The observational and reanalysis data obtained to perform the analysis as well as the methodology

followed to construct reanalysis-derived datasets to be compared to observations are described in Section 2. The results are presented in Section 3, which includes a comparison between reanalysis and observations, an analysis of reanalysis-derived seasonal variability, and analysis of interannual variability in capacity factors for hypothetical wind farms at the location of the observing stations. Finally, the conclusions are given in Section 4.

## 2 Data and methodology

In this study, wind speed observations from a number of locations across Mexico will be compared with wind data obtained from the reanalysis data of the Modern-Era Retrospective Analysis for Research Applications version 2 (MERRA-2) [15]. This section provides an overview of the datasets used.

### 2.1 Wind speed Observations

Measurement stations at 24 widespread locations in the Mexican territory were used in this study. The datasets were recorded with a 10-minute time resolution. Data availability for each station for the period 2004 to 2008 is shown in figure 1. Each of the stations have different measurement levels ranging between 10 and 80 m above ground level.

### 2.2 Wind farm data

Lack of information in Mexico not only affects wind speed measurements, but also the wind power production per wind farm. The only freely available reported data about wind farm production in Mexico is published by the Energy Secretariat, SENER [33, 34, 35, 36]. These reports are published since 2015 as part of the perspectives and plans for the electric sector development. Reported data has information related to each wind farm’s installed capacity and annual production, enough to calculate the annual Capacity Factor (CF) at each location.

For one wind turbine (Vestas V90-2 MW) at a wind farm in Arriaga, Chiapas data was obtained for the whole of 2016. Collected data includes

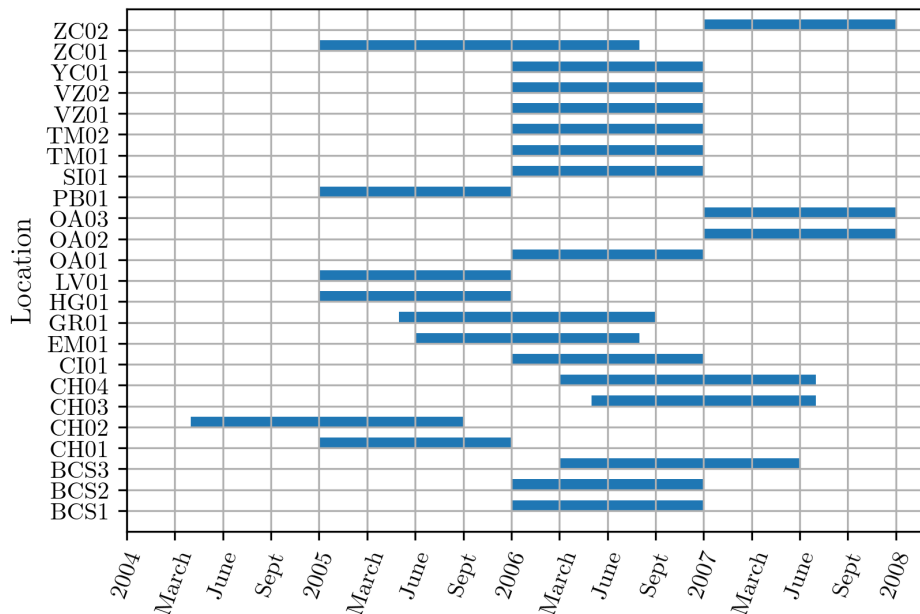


Figure 1: Data availability for the measured data.

measurements of wind speed and power generation with a 10-minute time resolution. Wind measurements at this location were added to the comparison between MERRA-2 and the measurements as another station. The data was measured at 80 m above displacement height, which is Vestas V90-2 MW hub-height. General locations of wind speed masts and wind farms are presented in figure 2, names by region are presented in figure 3.

### 2.2.1 Reanalysis

MERRA-2 is a global reanalysis with an approximate resolution of  $0.5^\circ \times 0.625^\circ$  and 72 vertical levels from the surface to about 0.01 hPa [15]. Mexico is located between  $14^\circ\text{N}$  and  $33^\circ\text{N}$ , for which the resolution of the reanalysis corresponds to a grid spacing of about 50 km. The MERRA-2 data subset used is delimited between  $80^\circ$  and  $120^\circ$  West, and from  $10^\circ$  to  $35^\circ$  North, which includes the whole Mexican territory. This domain is displayed in figure 2. The variables included in the study are the northward and eastward

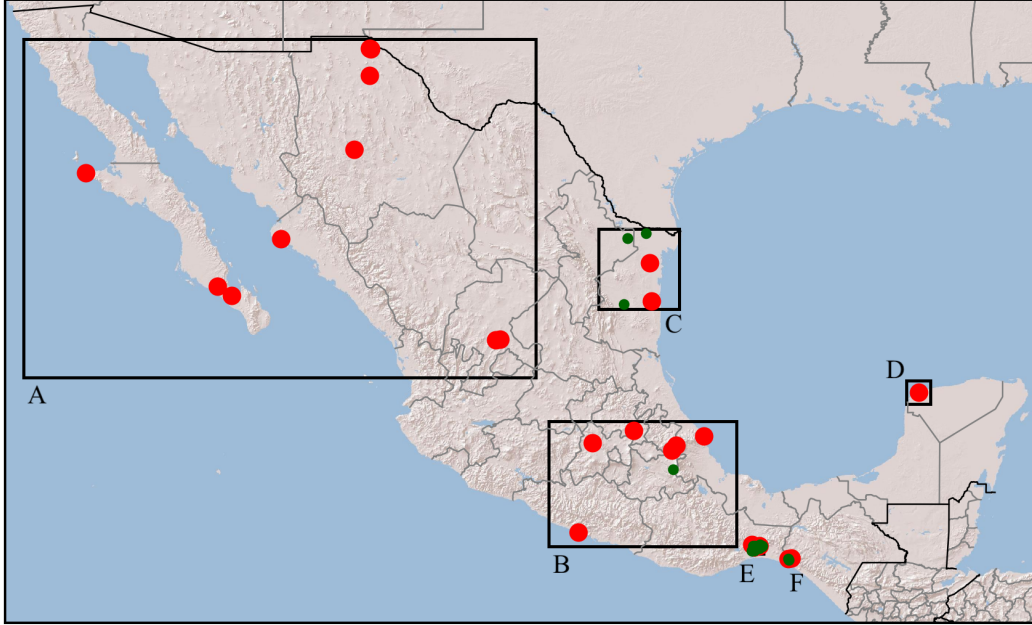
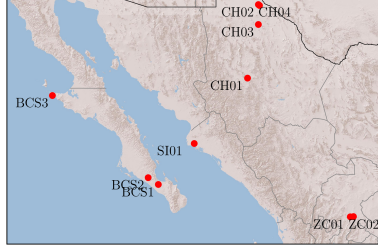


Figure 2: MERRA-2 used domain with the measurement stations in red. Wind farms marked in green.

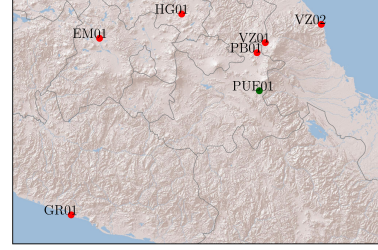
wind at 2, 10 and 50 m above displacement height. The MERRA-2 product used has a 1-hourly resolution therefore to make both MERRA-2 and measurement datasets comparable, the temporal resolution of the measurement data had to be decreased from 10 minutes to 1 hour. The data was obtained for the period from 1980 to 2018, inclusive.

Based on Cannon et al. [5], wind speed at the locations of the measurement station is calculated from MERRA-2 wind data with a bi-linear interpolation using the nearest four points of the MERRA-2 data grid for the three nominal MERRA-2 heights above ground level (2, 10 and 50 m). For each station, the MERRA-2-derived wind speeds are vertically interpolated to the specified height where the measurements were taken assuming a logarithmic vertical wind profile.

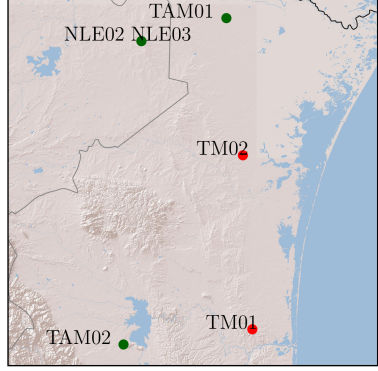
Reanalyses exhibit biases primarily related to the underlying weather forecast model, the inability to solve detailed topography and other sub-grid scale phenomena [38]. Although reanalyses represent a very useful tool for the characterisation of wind in a given region lacking observations, a calibration



(a) Northwest region.



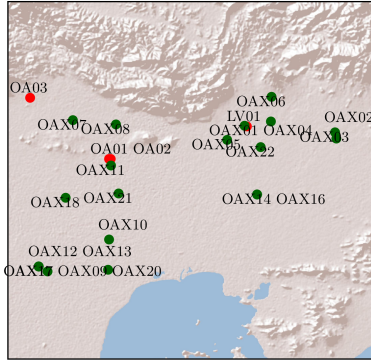
(b) Center region.



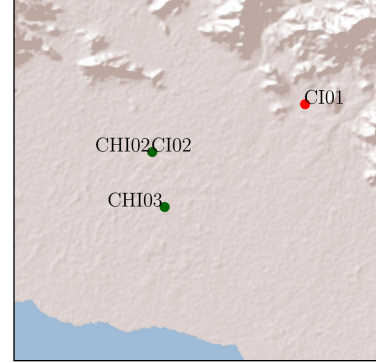
(c) Northeast region.



(d) Yucatan region.



(e) La Ventosa, Oaxaca.



(f) Chiapas.

Figure 3: Zoom-in for the different regions studied. Marked in red the anemometric stations and the associated names. In green the locations of the wind farms selected to analyze the reliability of MERRA-2 to estimate power production.

is necessary to achieve the best possible results. In this work, biases in the wind speed derived from MERRA-2 were corrected using a quantile mapping method at each site [27]. The quantile mappings were calibrated using the available observations for the available observational period and assumed valid for the rest of the MERRA-2 period (1980–2018).

The methodology followed to determine wind power data from MERRA-2 wind speeds is similar to the Virtual Wind Farm (VWF) described in [38, 37]. After bias-correcting the wind speed data, the next step in the VWF methodology requires the transformation of wind speeds into power production. According to the reported data of Mexican wind farms, three of the most common wind turbine models used in the country are the Acciona AW70-1.5 MW, Gamesa G80-2 MW and Vestas V90-2 MW. Hence, these models are used as a reference to calculate the wind power output produced at each one of the measurement locations. Since the V90-2MW is the model used in the Arriaga wind farm, for which we have measured wind speed and power output data, we use the data corresponding to this wind turbine to compute the power output at that site. The power curves of the three wind turbines are displayed in figure 4. The hub-height used is determined by the most common tower height in Mexico for each wind turbine: 80 m for the AW70-1.5 MW, 67 m for the G80-2 MW, and 80 m for the V90-2 MW. Hourly CF is estimated with the power curves and the reanalysis-derived wind speed time series.

## 3 Results

### 3.1 Wind speed correlation coefficient analysis

In this section, we present a comparison between hourly-averaged observations and the hourly wind speeds estimated by MERRA-2. The Pearson correlation coefficient,  $r$ , was used to measure the degree of linear correlation between these two datasets. Scatter plots with smoothed densities [12] of the bias-corrected hourly wind speeds derived from MERRA-2 against hourly-

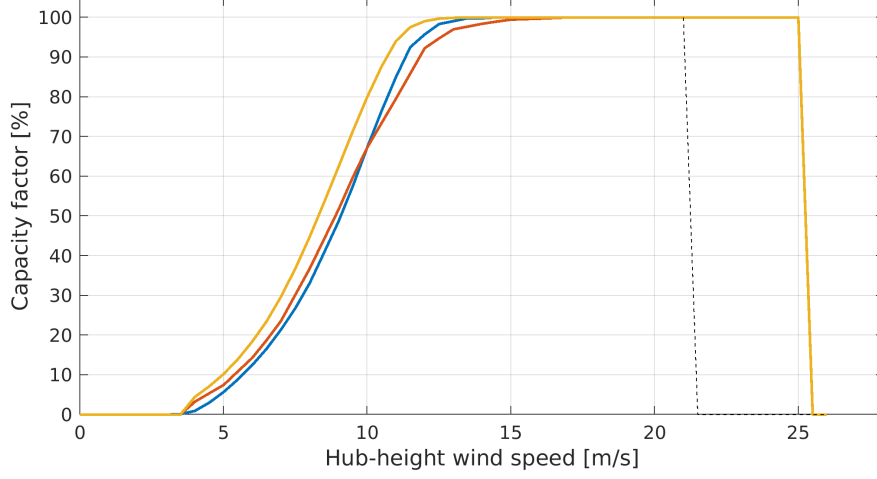


Figure 4: Power curves used in the analysis: Acciona AW70-1.5 MW in blue, Gamesa G80-2 MW in red and Vestas V90-2 MW in yellow. The black dotted line shows the reconnection cut-in wind speed.

averaged observations at BCS3, YC01 and OA02 are shown in Figure 5. BCS3 is the station that exhibits the lowest correlation coefficient ( $r = 0.38$ , Figure 5a). This station is located in the Baja California Peninsula, which separates the Gulf of California from the Pacific Ocean. The peninsula is about 40 km at its narrowest point and only a few hundred kilometres at its widest point. Furthermore, mountain ranges run along the peninsula, with a highest point about 3 km a.s.l. This complex topography and the air-land and air-ocean interactions that take place around the peninsula are difficult to represent accurately in a numerical model with a grid spacing of around 50 km at the latitudes at which Mexico is located. Therefore, it is possibly not surprising to find that the correlation between the global reanalysis and punctual observations at this site are not well correlated. Under these conditions the quantile mapping bias correction has a small effect (Figure 5b).

YC01 is located in the Yucatán Peninsula. In contrast with the Baja California Peninsula, the Yucatán Peninsula is a wide mostly flat region and therefore even though the winds at YC01 are determined by sea-land contrasts and their interaction with the atmosphere, just as are those over



Baja California, the large-scale circulation plays a more important role. The correlation coefficient between the observations and the reanalysis at this station is  $r = 0.70$ , and while the uncorrected wind speeds derived from MERRA-2 underestimate the observations, especially for large wind-speed values, applying bias correction improves this situation (Figure 5b).

OA02 is located in the region known as La Ventosa, which concentrates about 60% of Mexico’s wind power installed capacity. La Ventosa is located within the Chivela Pass [39], which is a gap between two mountain ranges, namely the Sierra Madre Oriental to the west and the Sierra Madre de Chiapas to the east. This particular topography induces episodes of very strong gap-outflow wind over the Gulf of Tehuantepec [39], when the large-scale forcing is favourable for the development of northerly wind due to the occurrence of, for example, Central American cold surges [30, 31] (or *Nortes* as they are known in Mexico). The annual wind probability distribution function in the region is characterised by a bimodal distribution [20]. This feature is present both in the available observations and in the reanalysis-derived data (Figure 5c). The correlation coefficient between the observations and the reanalysis at this station is  $r = 0.84$ , one of the highest for hourly data. When uncorrected, MERRA-2 slightly overestimates the observed data at low wind speeds, and underestimates it at higher wind speed values. As in the case of YC01, the effect of bias correction in correcting this mismatch is noticeable (Figure 5c).

It is clear by looking at this three-station sample (Figure 5) that the hourly data is better represented by the reanalysis at some sites than at others. A range of  $r$  values between 0.38 and 0.88 was found amongst the stations, which indicates that the reanalysis’ performance is not uniform throughout the country (See *Hourly* column in Table 1). The highest correlation between the reanalysis and observations was found at stations located in the Chivela Pass, including the stations located in La Ventosa region (OA02 and OA03) and Chiapas (CI01 and CI02). These stations exhibit correlation coefficients above 0.80 at hourly resolution (Table 1). However, in other regions the correlation coefficient can be very low. In particular, stations in the Baja California Peninsula (BCS1, BCS2 and BC3) exhibit correla-

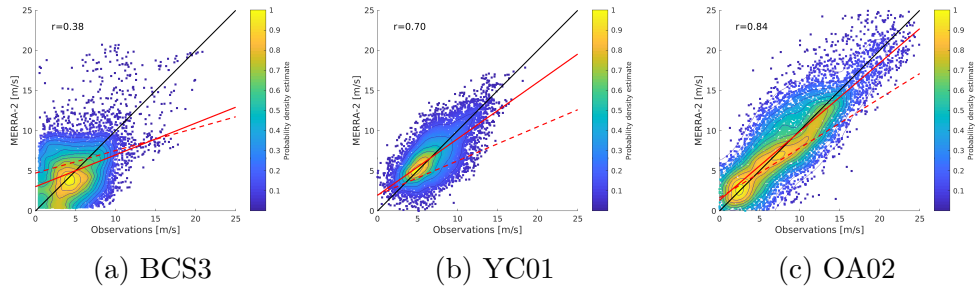


Figure 5: Scatter plots of 40-m wind speed observations against MERRA-2 for (a) BCS3, (b) YC01 and (b) OA02. The shading indicates the density of points in the neighbourhood. The red lines represent the linear regression of MERRA-2 without bias correction (dashed) and with bias correction (solid). The black line represents the identity line. The Pearson correlation coefficient,  $r$ , is indicated in the top left corner in each panel.

tion coefficients below 0.50 at the same hourly resolution (Table 1), which is consistent with the discussion above. Other stations that also exhibit low correlation at hourly resolution are those located in the central region (GR01, EM01, HG01, PB01 and VZ01) (Table 1). This region is characterised by the presence of the Trans-Mexican Volcanic Belt, and other high-topography features, which are difficult to capture accurately in a low resolution global reanalysis.

The correlation between hourly punctual observations and the low resolution reanalysis is largely limited by the inaccurate representation of local effects, which ultimately determine the wind characteristics at a given site. However, we hypothesise that the influence of local effects would decrease and the effects of the driving large-scale circulation would increase in a signal averaged over longer time scales. A further analysis has therefore been completed to determine how the correlation coefficient of the data changes when the time average is increased from hourly to weekly. Each station’s correlation coefficients as a function of time average are shown in Table 1. As expected, the correlation coefficient increases as the time average increases. Thus, while there are only 10 stations with correlation coefficients above 0.7 at hourly resolution, there are 18 stations reaching the same level of accuracy at daily and even 19 stations at weekly resolution. The increase of correlation

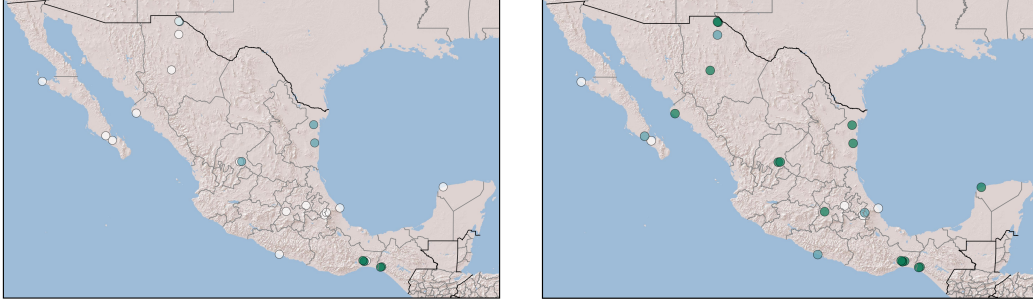
Region	Station	Altitude (m.a.s.l.)	Hourly	3-Hourly	6-Hourly	12-Hourly	Daily	Weekly
BC	BCS1	271	0.40	0.42	0.44	0.51	0.54	0.53
	BCS2	142	0.43	0.47	0.51	0.61	0.71	0.77
	BCS3	17	0.38	0.40	0.43	0.50	0.56	0.59
NW	CH01	2080	0.66	0.70	0.74	0.77	0.85	0.89
	CH02	1499	0.71	0.75	0.78	0.81	0.84	0.84
	CH03	1221	0.57	0.62	0.66	0.69	0.76	0.71
	CH04	1287	0.69	0.73	0.76	0.81	0.84	0.88
	SI01	11	0.63	0.67	0.72	0.77	0.81	0.89
NC	ZC01	2050	0.64	0.68	0.72	0.76	0.79	0.87
	ZC02	2714	0.71	0.74	0.77	0.82	0.88	0.92
NE	TM01	25	0.79	0.81	0.83	0.86	0.89	0.90
	TM02	43	0.78	0.80	0.82	0.85	0.88	0.90
SW	GR01	0	0.54	0.58	0.61	0.64	0.59	0.72
TVB	EM01	2551	0.55	0.60	0.65	0.68	0.78	0.84
	HG01	2811	0.44	0.46	0.48	0.53	0.46	0.47
	PB01	2332	0.40	0.42	0.44	0.47	0.56	0.67
	VZ01	2398	0.54	0.57	0.60	0.62	0.64	0.70
	VZ02	17	0.62	0.64	0.67	0.69	0.69	0.60
YC	YC01	0	0.70	0.73	0.76	0.87	0.92	0.95
CP	LV01	70	0.69	0.70	0.72	0.74	0.76	0.80
	OA01	30	0.79	0.80	0.82	0.84	0.87	0.90
	OA02	30	0.85	0.86	0.88	0.90	0.92	0.93
	OA03	70	0.82	0.84	0.86	0.89	0.91	0.94
	CI01	50	0.82	0.84	0.86	0.88	0.91	0.95
	CI02	90	0.88	0.89	0.90	0.92	0.93	0.97

Table 1: Correlation coefficients between the observations at the maximum height at each measuring station and MERRA-2 after bias correction for a selection of time averages ranging from hourly to weekly. The regions acronyms stand for Baja California (BC), North-West (NW), North-Centre (NC), North-East (NE), South-West (SW), Trans-Mexican Volcanic Belt (TVB), Yucatán (YC) and Chivela Pass (CP). The cells are colour-coded as follows: high correlation (green), medium correlation (light blue) and low correlation (not highlighted).

with time average produces distinctive regional patterns across the country, which has led us to classify the stations in eight regions: The Baja California Peninsula, the North-West, the North-Centre, the North-East, the South-West, the Trans-Mexican Volcanic Belt, Yucatán and the Chivela Pass. The Trans-Mexican Volcanic Belt includes the two stations in Veracruz (VZ01 and VZ02) due to their proximity to and latitudinal alignment with this orographic feature even though they do not strictly form part of it. These regions are shown in Figure 6, in which the stations are colour-coded according to their correlation coefficients at hourly and weekly resolutions.

The stations in the Chivela Pass region are clearly the most accurately represented by MERRA-2. Even though there are two stations (LV01 and OA01) which have hourly correlations below 0.8, the 6-hourly correlation for all but one of the stations in this region is above this threshold. Similarly, the stations in the North-East exhibit high correlation coefficients above 0.7 for all time averages and above 0.8 at 6-hourly resolution. By contrast, the stations in the Baja California Peninsula, the South-West and the Trans-Mexican Volcanic Belt generally exhibit low correlations even for long time averages. Thus, even at a daily resolution the correlation coefficients are below 0.7 for all but two stations in these regions. The stations in the remaining regions (North-West, North-Centre and Yucatán) exhibit some variability in the correlation coefficient at a given time average, but in general their 6-hourly correlations are above 0.7 and their daily correlations are above 0.8.

From these results we conclude that MERRA-2 struggles to represent the surface flow around topographic features of high spatial variability such as mountain ranges (e.g. stations in the Trans-Mexican Volcanic Belt), rapid transition between land and sea (e.g. stations in the Baja California Peninsula) or high topography close to ocean (GR01 and VZ01 and VZ02). The flow near the surface around all these features crucially depends on the local surface characteristics (e.g. terrain, sub-grid terrain variability, roughness length), which determine the fluxes and, ultimately, the wind within the boundary layer. On the other hand, the reanalysis can represent well near-surface flow in regions where it is determined by the large-scale circulation and its interaction with the large-scale orography. This is the case of the stations located within the Chivela Pass in which the occurrence of outbreaks of cold air masses moving south from the North American Great Plains [30] lead to a rise in pressure over the Gulf of Mexico. This creates a pressure gradient along the Chivela Pass and leads to the development of strong northerly *gap winds* across the Pass [26, 30, 39]. This is also the case of Yucatán and the stations in the North-East region, whose wind speed characteristics are dominated by the influence of the trade winds and their deflection to the north by the high orography of the Sierra Madre Oriental to form the western flank of the North Atlantic subtropical anticyclone [10, 41].



(a) Correlation coefficients of the time series at *hourly* resolution.

(b) Correlation coefficients of the time series at *weekly* resolution.

Figure 6: Locations of the twenty-five stations for (a) hourly and (b) weekly time averaging. The markers colours represent the value of the correlation coefficient as follows:  $r < 0.7$  (white),  $0.7 \leq r < 0.8$  (light blue) and  $r \geq 0.8$  (green).

Another aspect in this analysis is that there are stations separated from each other by a few hundreds of kilometres with no apparent difference in the general features of the surrounding topography, whose reanalysis representations exhibit nevertheless very different accuracy. This is the case, for example, of CH03, which displays a correlation coefficient below 0.7 even at a time average of 12 hours in contrast with three correlations above 0.8 of CH02 and CH04, located just about 100 km to the north. As a result, excluding the North-East and those regions represented by a single station (i.e. South-West and Yucatán), the rest of the regions are not completely homogeneous in the way the correlation coefficient changes as the time average increases. Thus, even though a region might appear very well represented by the reanalysis caution must be exercised when using the reanalysis instead of observations for wind resource assessment.

### 3.2 Capacity factors analysis

The previous section of this paper analysed how well the MERRA2 reanalysis represents the variability of wind speed at a range of sites in Mexico. The aim of this section is to determine whether the reanalysis data can be used to estimate the wind power characteristics at the same sites. Specifically, the

analysis will focus on the annual capacity factor which is a key parameter in the economic assessment of wind farms.

The Mexican Wind Power industry annually reports wind farm power production along with information about the location of wind farms and the specific technology installed. From this information, the capacity factor of 27 wind farms has been calculated for the period 2014 to 2017. The measured values have been compared to values derived by applying the appropriate wind turbine power curve to the MERRA-2 wind speeds. The comparison has been completed on a regional resolution, whereby all wind farms were allocated to the observation location within 100 km radius. Any differences between the MERRA-2 derived capacity factors and the measured values are therefore due to either (1) errors in the wind speed or (2) the conversion from wind speed to power.

To determine the errors associated with the wind speeds, at three sites the capacity factor of the Vestas 2 MW turbine has been estimated by applying the power curve to the wind speed observations and the reanalysis wind speeds. Table 2 shows that for all three sites the derived annual capacity factor using the measured wind speeds is very similar to that from the MERRA-2 data. This indicates the error associated with the representation of the wind speed is relatively small. However, for CI02 (the one site for which there is both wind speed observations and a reported capacity factor), Table 2 shows that observed capacity factor is considerably lower: 0.34 in comparison to 0.39 for the observed wind speeds and 0.40 for the MERRA-2 wind speeds.

Figure 7 shows the comparison between the MERRA-2 derived annual capacity factor for that region and the observed capacity factors of all of the wind farms in that region [33, 34, 35, 36]. The multiple MERRA-2 values indicate the different years associated with the observed values. For all three regions, there is a wide dispersion in the observed performance of the wind farms. However, most of these values are above 20%, which is commonly associated as a lower limit production for economic feasibility. The variation of the performance between farms within the same region are associated with either differences in the technology and operations between sites and/or local

Station	Year	CF measured	CF from stations	CF from MERRA-2
CI02@80m	2016	0.34	0.39	0.40
LV01@32m	2005	-	0.61	0.61
OA02@80m	2007	-	0.54	0.53

Table 2: Capacity factors from three anemometric stations at La Ventosa region. CF’s from stations and MERRA-2 presents similar values; between these values, the maximum difference is one percent. Real CF is lower; this may be explained because the method considers continuous operation along the year.

variations in the wind resource.

It is also clear from Figure 7 that the observed capacity factors are considerably lower than the MERRA-2 derived values. This suggests that the power curve conversion method is too simplistic and leads to an overestimate of the generation. To derive a more accurate value of the output of a wind farm, other factors need to be considered such as the wind farm operations and turbine downtime and maintenance. There are two factors that might be leading to this bias. The first would be the implicit assumptions in the power curve conversion method, namely that the wind is always aligned with the wind turbine axis and that power only depends on wind speed. The second is the lack of information about the wind farms’ development (e.g. new installed capacity) and operation (e.g. maintenance stops). Therefore, to derive a more accurate values of the output of a wind farm, it is essential not only to take a more complex approach in the conversion from wind speed and direction into wind power production but also to promote the creation of Mexican data, including annual power production and Operation and Maintenance stops.

Figure 8 shows the variability of the annual capacity factor at each site derived from the reanalysis across the 39 years. Based on the analysis described above, the magnitude of the capacity factors derived using the reanalysis should be considered as an upper bound (i.e. if the turbine was operating according to the power curve throughout the whole year). While the magnitudes might not be truly reflective of observed output, an understanding

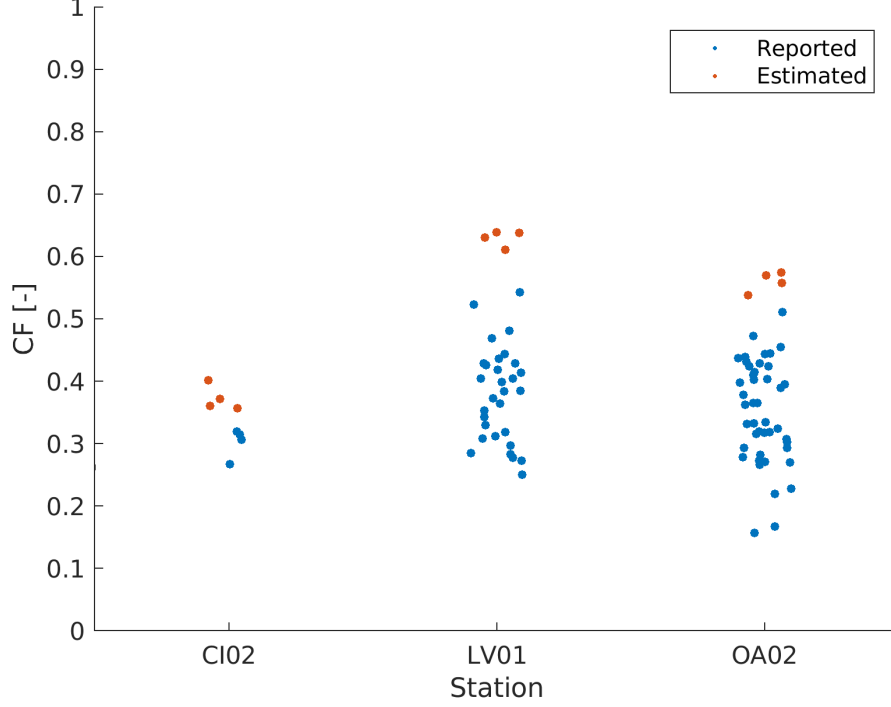


Figure 7: Annual CF based on the MERRA-2 bias-corrected data at measurement station with a wind turbine V90-2 MW in red dots, and annual CF reported by SENER for the wind farms within 100 km of the measurement station, in blue dots.

of the possible year to year variability could be valuable when considering the management of a wind farm and power system. For each station the boxplots represent the average annual CF with a box of the 25th and 75th percentiles and the line extended to the most extreme values. The analysis shows large levels of inter-annual variability in the wind generation at each location. For example, in CH01, the annual capacity factor can vary from 0.18 to 0.30. This can have significant implications for the annual revenue from the wind farm and should be considered in an economic assessment. Furthermore, sufficient levels of alternative capacity need to be procured to protect against a low wind yield year.



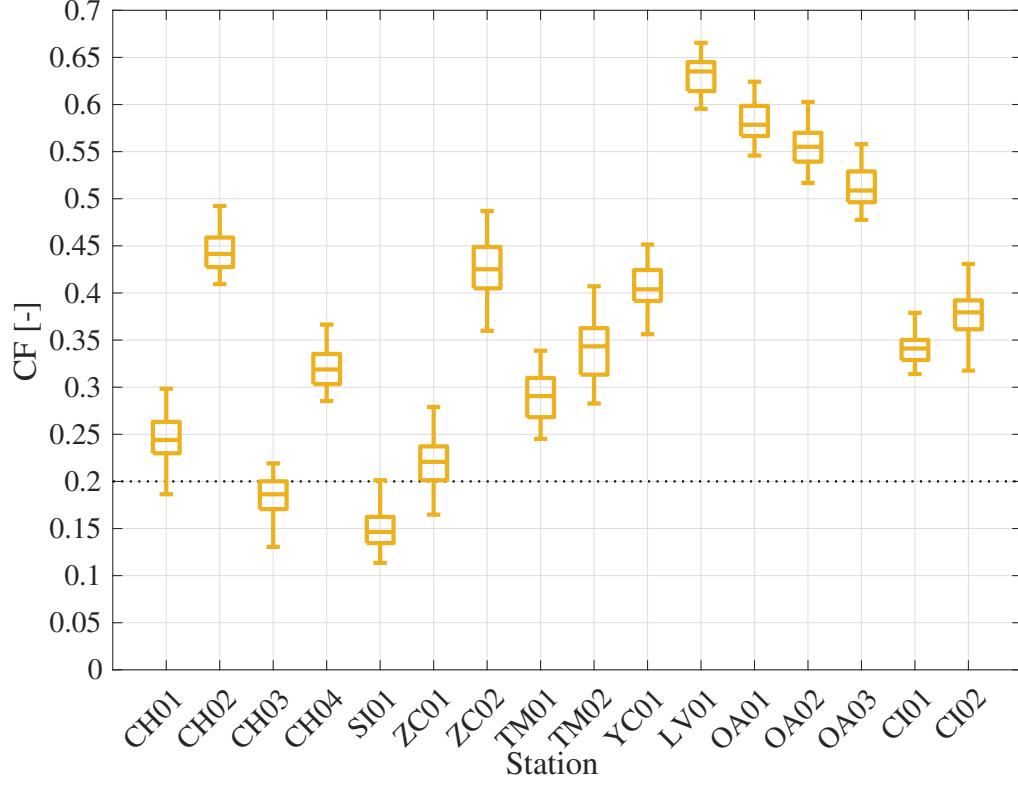


Figure 8: Capacity factors at each location are presented for the wind turbine V90-2 MW. Similar behavior was observed for all wind turbines; however, it is only showed the best-performed among the three technologies assessed. The black dotted line marks the CF 0.2 threshold.

## 4 Conclusions

In this work we have assessed the ability of the global reanalysis MERRA-2 to represent near-surface wind speeds over Mexico taking as reference 24 measuring stations distributed throughout the country. The assessment was carried out using the Pearson correlation coefficient between the observed time series and those obtained by interpolating horizontally and vertically from the reanalysis grid to the location of the observing towers. The reanalysis-derived winds were then used as an input to calculate the interannual variability in the performance of hypothetical wind turbines located at the observing tower

positions.

The results of the analysis indicate that the performance of MERRA-2 at representing wind speed over Mexico is not uniform throughout the country. Based on geographical location and the performance of MERRA-2, the stations were grouped into eight regions. The best-represented region was the Chivela Pass, which included stations in Oaxaca and Chiapas near the Gulf of Tehuantepec where the wind is intense throughout the year. It is noticed that regions where the largest discrepancies were found were those in which the land-sea mask (Baja California) or the orography (South-West and the Trans-Mexican Volcanic Belt) might not be sufficiently accurate due to the reanalysis resolution.

Both observations and MERRA-2-estimated wind speed were time-average over periods ranging from one hour to one week. The performance generally improves as the time-averaging period increases, so that the annual cycle is well represented by the reanalysis in almost all stations, although there are some stations that show no or little improvement as the time-averaging period increases.

After characterising the reanalysis' skill to represent wind speed variability, we assessed capacity factor variability for those regions in which wind speed variability is found to be well-represented. The availability of one year's worth of measurements of wind speed and capacity factor for a wind farm in Chiapas enabled the assessment of the error involved in using MERRA-2-estimated wind speed. This error was found to be small (about 3%). However, it was also found that the capacity factor computed with MERRA-2 wind speed produce an overestimation (of about 18% for the measured year at the wind farm in Chiapas) in the magnitude of the actual power output reported by wind farms, and an underestimation of variability. This may be attributable to the relative simplicity of the power curve of the turbine that was used in the analysis, as this does not represent operational details, such as maintenance downtime, of the actual turbines V90-2 MW installed in Chiapas. However, more work and data would be needed to address this point in a more comprehensive manner. To acknowledge this caveat the MERRA-2-estimated capacity factors were interpreted as upper bounds for the actual

capacity factors.

Annual capacity factors for hypothetical wind farms at the locations of the observing towers were computed for the 39 years in the MERRA-2 dataset. The results show large annual capacity factor interannual variability at several sites. For some, annual capacity factor would fluctuate in such a way that some years would lie below the lower limit production for economic feasibility. This highlights the importance of considering interannual variability in the assessment of wind resource and the need for accurate seasonal forecasting systems to plan ahead of a low wind yield season.

These results show that while reanalyses can be a valuable tool for the assessment of interannual variability and other long-term aspects of wind power production, it is necessary to be cautious. We have shown that a reanalysis performance largely depends on the target region, and post-processing via e.g. bias correction towards an observational dataset is usually desirable. However, while bias correction can be applied to improve the match between reanalysis estimations and observations, the bias correction is only strictly valid for the observation location. Furthermore, in some locations bias correction is not sufficient to bring the reanalysis to an acceptable accuracy level. Once these caveats have been considered, wind power output estimations based on atmospheric reanalyses can become a valuable tool to understand the feasibility and installed capacity requirements for Mexico to meet its renewable energy targets.

## Acknowledgements

The authors would like to thank Eng. Juan Ramón López Samayoa, Manager of the Wind Farm for his support in the development of this research. The authors would like to thank Eng. Ricardo Saldaña Flores and MSc. Ubaldo Miranda Miranda through the project “Plan de acción para eliminar barreras para la implantación comercial de la generación eoloelectrica en México” for the support in the development of this research.

## Funding

This research was partially funded by:

**Mexico CONACyT - SENER - Sustentabilidad Energética, Project 272063**, “Strengthening of the field of Wind Energy in the Doctoral Program in Engineering Field of Knowledge in Energy based in the Institute of Renewable Energies of the National Autonomous University of Mexico”; (*Institutional Strengthening for Energy Sustainability*).

**Newton Fund Institutional Links grant, ID 432335407**, funded by the UK Department of Business, Energy and Industrial Strategy and delivered by the British Council, and by the UK National Centre for Atmospheric Science through the Atmospheric hazard in developing Countries: Risk assessment and Early Warning (ACREW) NE/R000034/1 project.

**Programa de Becas Estudiantiles SEP-UNAM-FUNAM 2017 (Beca de Capacitación en Métodos de Investigación – Primera Fase)**. Through this grant, Carlos Francisco Morales Ruvalcaba developed a short research stay at the University of Reading, UK.

## References

- [1] Alexandri, R., Muñoz, L., Guerrero, L., Rodríguez, F., Ubaldo, A., Ramos, A., 2016. Prospectiva de energías renovables 2016-2030. Technical Report. Secretaría de Energía, México.
- [2] Beltran, L., Villanueva Arcos, E., Muñozcano, L., Rodriguez, J., Ramirez, M., Gasca, S., Ramírez, A., Rocha, D., Avila, D., Navarrete, E., Rangel, R., Carrasco, F., 2017. Reporte de avance de energías limpias. Technical Report. Secretaría de Energía.
- [3] Benedek, J., Sebestyén, T.T., Bartók, B., 2018. Evaluation of renewable energy sources in peripheral areas and renewable energy-based rural development. Renewable and Sustainable Energy Reviews

- 90, 516–535. URL: <https://doi.org/10.1016/j.rser.2018.03.020>, doi:10.1016/j.rser.2018.03.020.
- [4] Bosch, J., Staffell, I., Hawkes, A.D., 2018. Temporally explicit and spatially resolved global offshore wind energy potentials. *Energy* 163, 766–781. URL: <https://doi.org/10.1016/j.energy.2018.08.153>, doi:10.1016/j.energy.2018.08.153.
- [5] Cannon, D.J., Brayshaw, D.J., Methven, J., Coker, P.J., Lenaghan, D., 2015. Using reanalysis data to quantify extreme wind power generation statistics: A 33 year case study in Great Britain. *Renewable Energy* 75, 767–778. URL: <http://dx.doi.org/10.1016/j.renene.2014.10.024> papers3://publication/doi/10.1016/j.renene.2014.10.024.
- [6] Carrasco-Díaz, M., Rivas, D., Orozco-Contreras, M., Sánchez-Montante, O., 2015. An assessment of wind power potential along the coast of Tamaulipas, northeastern Mexico. *Renewable Energy* 78, 295–305. doi:10.1016/j.renene.2015.01.007.
- [7] Chen, W., Castruccio, S., Genton, M.G., Crippa, P., 2018. Current and Future Estimates of Wind Energy Potential Over Saudi Arabia. *Journal of Geophysical Research: Atmospheres* 123, 6443–6459. doi:10.1029/2017JD028212.
- [8] Collins, S., Deane, P., Ó Gallachóir, B., Pfenninger, S., Staffell, I., 2018. Impacts of Inter-annual Wind and Solar Variations on the European Power System. *Joule* 2, 2076–2090. doi:10.1016/j.joule.2018.06.020.
- [9] Cradden, L.C., McDermott, F., Zubiate, L., Sweeney, C., O’Malley, M., 2017. A 34-year simulation of wind generation potential for Ireland and the impact of large-scale atmospheric pressure patterns. *Renewable Energy* 106, 165–176. URL: <http://dx.doi.org/10.1016/j.renene.2016.12.079>, doi:10.1016/j.renene.2016.12.079.
- [10] Davis, R.E., Hayden, B.P., Gay, D.A., Phillips, W.L., Jones, G.V., 1997. The North Atlantic subtropical anticyclone. *J. Clim.* 10, 728–744.

- [11] Dolter, B., Rivers, N., 2018. The cost of decarbonizing the canadian electricity system. *Energy Policy* 113, 135 – 148. URL: <http://www.sciencedirect.com/science/article/pii/S0301421517307140>, doi:<https://doi.org/10.1016/j.enpol.2017.10.040>.
- [12] Eilers, P., Goeman, J.J., 2004. Enhancing scatterplots with smoothed densities. *Bioinformatics* 20, 623–U82. doi:[10.1093/bioinformatics/btg454](https://doi.org/10.1093/bioinformatics/btg454).
- [13] Figueroa-Espinoza, B., Salles, P., Zavala-Hidalgo, J., 2014. On the wind power potential in the northwest of the Yucatan Peninsula in Mexico. *Atmósfera* .
- [14] Gao, M., Ding, Y., Song, S., Lu, X., Chen, X., McElroy, M.B., 2018. Secular decrease of wind power potential in India associated with warming in the Indian Ocean. *Science Advances* 4, eaat5256. URL: <http://advances.sciencemag.org/content/4/12/eaat5256.full>, doi:[10.1126/sciadv.aat5256](https://doi.org/10.1126/sciadv.aat5256).
- [15] Gelaro, R., McCarty, W., Suárez, M.J., Todling, R., Molod, A., Takacs, L., Randles, C.A., Darmenov, A., Bosilovich, M.G., Reichle, R., Wargan, K., Coy, L., Cullather, R., Draper, C., Akella, S., Buchard, V., Conaty, A., da Silva, A.M., Gu, W., Kim, G.K., Koster, R., Lucchesi, R., Merkova, D., Nielsen, J.E., Partyka, G., Pawson, S., Putman, W., Rienecker, M., Schubert, S.D., Sienkiewicz, M., Zhao, B., 2017. The modern-era retrospective analysis for research and applications, version 2 (MERRA-2). *Journal of Climate* 30, 5419–5454. doi:[10.1175/JCLI-D-16-0758.1](https://doi.org/10.1175/JCLI-D-16-0758.1).
- [16] Global Wind Energy Council, 2018. GLOBAL WIND REPORT. Technical Report. Global Wind Energy Council.
- [17] Hernandez, A., Peña, R., Mendez, W., Visairo, N., Nunez, C., 2013. Wind resource assessment in the surroundings of San Luis Potosí, Mexico, in: 2013 IEEE International Autumn Meeting on Power, Electronics and Computing, ROPEC 2013, pp. 1–6.

- [18] Hernandez-Escobedo, Q., 2016. Wind Energy Assessment for Small Urban Communities in the Baja California Peninsula, Mexico. *Energies* 9, 805. doi:10.3390/en9100805.
- [19] Hernández-Escobedo, Q., Saldaña-Flores, R., Rodríguez-García, E.R., Manzano-Agugliaro, F., 2014. Wind energy resource in Northern Mexico. *Renewable and Sustainable Energy Reviews* 32, 890–914. doi:10.1016/j.rser.2014.01.043.
- [20] Jaramillo, O.A., Borja, M.A., 2004. Wind speed analysis in La Ventosa, Mexico: a bimodal probability distribution case. *Renewable Energy* .
- [21] Kim, J.Y., Oh, K.Y., Kim, M.S., Kim, K.Y., 2018. Evaluation and characterization of offshore wind resources with long-term met mast data corrected by wind lidar. *Renewable Energy* URL: <https://doi.org/10.1016/j.renene.2018.06.097>, doi:10.1016/j.renene.2018.06.097.
- [22] Mills, A.D., Millstein, D., Jeong, S., Lavin, L., Wiser, R., Bolinger, M., 2018. Estimating the value of offshore wind along the United States’ Eastern Coast. *Environmental Research Letters* 13. doi:10.1088/1748-9326/aada62.
- [23] Millstein, D., Solomon-Culp, J., Wang, M., Ullrich, P., Collier, C., 2018. Wind energy variability and links to regional and synoptic scale weather. *Climate Dynamics* 0, 0. URL: <http://dx.doi.org/10.1007/s00382-018-4421-y>, doi:10.1007/s00382-018-4421-y.
- [24] Olauson, J., 2018. ERA5: The new champion of wind power modelling? *Renewable Energy* 126, 322–331. URL: <https://doi.org/10.1016/j.renene.2018.03.056>, doi:10.1016/j.renene.2018.03.056.
- [25] Olauson, J., Bergkvist, M., 2015. Modelling the Swedish wind power production using MERRA reanalysis data. *Renewable Energy* 76,

- 717–725. URL: <http://dx.doi.org/10.1016/j.renene.2014.11.085>, doi:10.1016/j.renene.2014.11.085.
- [26] Overland, J.E., Walter Jr, B.A., 1981. Gap winds in the Strait of Juan de Fuca. *Mon. Weather Rev.* 109, 2221–2233.
- [27] Piani, C., Haerter, J.O., Coppola, E., 2010. Statistical bias correction for daily precipitation in regional climate models over Europe. *Theoretical and applied climatology*. 99, 187–192. doi:10.1007/s00704-009-0134-9.
- [28] Rienecker, M.M., Suarez, M.J., Gelaro, R., Todling, R., Bacmeister, J., Liu, E., Bosilovich, M.G., Schubert, S.D., Takacs, L., Kim, G.K., Bloom, S., Chen, J., Collins, D., Conaty, A., da Silva, A., Gu, W., Joiner, J., Koster, R.D., Lucchesi, R., Molod, A., Owens, T., Pawson, S., Pegion, P., Redder, C.R., Reichle, R., Robertson, F.R., Ruddick, A.G., Sienkiewicz, M., Woollen, J., 2011. MERRA: NASA’s Modern-Era Retrospective Analysis for Research and Applications. *J. Climate* .
- [29] Rose, S., Apt, J., 2015. What can reanalysis data tell us about wind power? *Renewable Energy* 83, 963–969. URL: <http://dx.doi.org/10.1016/j.renene.2015.05.027>, doi:10.1016/j.renene.2015.05.027.
- [30] Schultz, D.M., Bosart, L., Bracken, W.E., 1997. The 1993 Superstorm Cold Surge : Frontal Structure , Gap Flow , and Tropical Impact clogenesis east of the Rocky or Sierra Madre Mountains. *Mon. Weather Rev.* 1, 5–40.
- [31] Schultz, D.M., Bracken, W.E., Bosart, L.F., 1998. Planetary- and synoptic-scale signatures associated with Central American cold surges. *Mon. Weather Rev.* 126, 5–27. doi:10.1175/1520-0493(1998)126<0005:PASSSA>2.0.CO;2.
- [32] Schwartz, M.N., Elliott, D.L., 1995. Mexico Wind Resource Assessment Project. Technical Report. NREL. Golden, Colorado 80401.



- [33] Secretaría de Energía, 2015. Programa de Desarrollo del Sistema Eléctrico Nacional 2015-2029. Technical Report. Secretaría de Energía.
- [34] Secretaría de Energía, 2016. Programa de Desarrollo del Sistema Eléctrico Nacional 2016-2030. Technical Report. Secretaría de Energía.
- [35] Secretaría de Energía, 2017. Programa de Desarrollo del Sistema Eléctrico Nacional 2017-2031. Technical Report. Secretaría de Energía.
- [36] Secretaría de Energía, 2018. Programa de Desarrollo del Sistema Eléctrico Nacional 2018-2032. Technical Report. Secretaría de Energía.
- [37] Staffell, I., Green, R., 2014. How does wind farm performance decline with age? Renewable Energy 66, 775–786. URL: <http://dx.doi.org/10.1016/j.renene.2013.10.041>, doi:10.1016/j.renene.2013.10.041, arXiv:arXiv:1006.4405v1.
- [38] Staffell, I., Pfenninger, S., 2016. Using bias-corrected reanalysis to simulate current and future wind power output. Energy 114, 1224--1239. URL: <http://dx.doi.org/10.1016/j.energy.2016.08.068>, doi:10.1016/j.energy.2016.08.068.
- [39] Steenburgh, W.J., Schultz, D.M., Colle, B.A., 1998. The structure and evolution of gap outflow over the Gulf of Tehuantepec, Mexico. Mon. Weather Rev. 126, 2673--2691. doi:10.1175/1520-0493(1998)126<2673:TSAEOG>2.0.CO;2.
- [40] Wang, M., Ullrich, P., Millstein, D., 2018. Datasets on hub-height wind speed comparisons for wind farms in California. Data in Brief 19, 214--221. URL: <http://dx.doi.org/10.1016/j.dib.2018.05.031>, doi:10.1016/j.dib.2018.05.031.
- [41] Zhu, J., Liang, X.Z., 2013. Impacts of the Bermuda high on regional climate and ozone over the United States. J. Clim. 26, 1018--1032.